

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 30, 2002, 10:08:33 ; Search time 2532 Seconds

(without alignments)  
17692.226 Million cell updates/sec

Sequence: 1 atgagcgttgtaagttgca.....gtacatcaagggtgtctaa 2766

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Labeling - first 45 summaries

Database: EST  
1: em\_estba:\*  
2: em\_esthu:\*  
3: em\_estin:\*  
4: em\_estlu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_luv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1          | 813.8 | 29.4        | 887    | 13    | BI913344    |
| 2          | 743.4 | 26.9        | 791    | 13    | BI523145    |
| 3          | 739.8 | 26.7        | 792    | 13    | BI522813    |
| 4          | 673.2 | 24.3        | 775    | 14    | BO770745    |
| 5          | 657   | 23.8        | 665    | 14    | BO189572    |
| 6          | 560.2 | 20.3        | 676    | 10    | BB280958    |

|   |    |       |      |      |    |           |                     |
|---|----|-------|------|------|----|-----------|---------------------|
| C | 7  | 553.6 | 20.0 | 1062 | 17 | CNS04DXR  | AL286344 Tetradon   |
|   | 8  | 516   | 18.7 | 527  | 12 | BS910325  | BS910325 602805921  |
|   | 9  | 416.4 | 15.1 | 718  | 13 | BI756778  | BI756778 603024449  |
|   | 10 | 407.8 | 14.7 | 475  | 12 | BF190598  | BF190598 237175 MA  |
|   | 11 | 407.4 | 14.7 | 499  | 17 | BH349372  | BH349372 CH230-32M  |
|   | 12 | 372.8 | 13.5 | 3063 | 11 | BC023215  | BC023215 Homo sapi  |
|   | 13 | 368.2 | 13.3 | 817  | 13 | BI551720  | BI551720 60197384   |
|   | 14 | 361.8 | 13.1 | 655  | 12 | BC619145  | BC619145 602781252  |
|   | 15 | 359.4 | 13.0 | 677  | 12 | BC933782  | BC933782 602390677  |
|   | 16 | 359   | 13.0 | 569  | 13 | BI906774  | BI906774 603064517  |
|   | 17 | 345   | 12.5 | 638  | 12 | BG772726  | BG772726 602720880  |
|   | 18 | 338.2 | 12.2 | 1020 | 17 | CNS03928  | AL265769 Tetradon   |
|   | 19 | 320   | 11.6 | 384  | 17 | AZ258968  | AZ258968 RPCI-23-1  |
|   | 20 | 319.8 | 11.6 | 550  | 9  | AA511690  | AA511690 V916A04.r  |
|   | 21 | 303   | 11.0 | 462  | 9  | AL712986  | AL712986 DXF2P6867  |
|   | 22 | 295.8 | 10.7 | 693  | 10 | BB648018  | BB648018 BB648018   |
|   | 23 | 294.2 | 10.6 | 551  | 17 | FR0052027 | AL688306 Fugu rubr  |
|   | 24 | 291.4 | 10.5 | 876  | 9  | AL666430  | AL666430 AL666430   |
|   | 25 | 270.4 | 9.8  | 590  | 14 | BO187487  | BO187487 UI-E-EJ1-  |
|   | 26 | 267.2 | 9.7  | 1098 | 17 | CNS04YRK  | AL309449 Tetradon   |
|   | 27 | 261   | 9.4  | 261  | 14 | BM930281  | BM930281 UI-E-EJ1-  |
|   | 28 | 258.8 | 9.4  | 1010 | 13 | BM547961  | BM547961 AGENCOURT  |
|   | 29 | 257.8 | 9.3  | 421  | 12 | BS554724  | BS554724 dac30d12.  |
|   | 30 | 253.8 | 9.2  | 486  | 12 | BF515586  | BF515586 UI-H-BW1-  |
|   | 31 | 240.6 | 8.7  | 1052 | 17 | CNS01EN3  | AL174504 Tetradon   |
|   | 32 | 237.8 | 8.6  | 514  | 12 | BC371210  | BC371210 Tetradon   |
|   | 33 | 237.6 | 8.6  | 753  | 13 | BI875890  | BI875890 fl64d02.y  |
|   | 34 | 233.8 | 8.5  | 668  | 14 | BO284787  | BO284787 f294e06.x  |
|   | 35 | 213.6 | 7.7  | 885  | 17 | CNS013JA  | AL102928 Drosophila |
|   | 36 | 208.4 | 7.5  | 689  | 14 | BO263135  | BO263135 f294e06.y  |
|   | 37 | 208.4 | 7.5  | 731  | 13 | BJ144538  | BJ144538 BV952826   |
|   | 38 | 206.2 | 7.5  | 455  | 10 | AV952826  | AV952826 AV952826   |
|   | 39 | 201.8 | 7.3  | 1090 | 17 | CNS05E11  | AL333770 Tetradon   |
|   | 40 | 200.8 | 7.3  | 863  | 17 | CNS03ONR  | AL253584 Tetradon   |
|   | 41 | 198.2 | 7.2  | 481  | 14 | BO339473  | BO339473 PM1-NN022  |
|   | 42 | 198.2 | 7.2  | 481  | 14 | BO339495  | BO339495 PM1-NN022  |
|   | 43 | 193.8 | 7.0  | 493  | 10 | AW346838  | AW346838 29591 MAR  |
|   | 44 | 189.2 | 6.8  | 544  | 9  | AA733833  | AA733833 v004f02.r  |
|   | 45 | 187.8 | 6.8  | 750  | 17 | CNS04MA3  | AL297156 Tetradon   |

## ALIGNMENTS

RESULT 1  
BI913344  
LOCUS 887 bp mRNA linear EST 16-OCT-2001  
DEFINITION 603178823F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5243308 5',  
mRNA sequence.  
ACCESSION BI913344  
VERSION BI913344.1 GI:16177710  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 887)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1613 row: m column: 05  
High quality sequence stop: 782.  
Location/Qualifiers 1. 887

FEATURES  
source



```

Oy 2018 CGGTGGACAACTGATCAAGAGAACAAACCTGGCCCTTGGTGTGGGGACCATTCCTCGGA 2077
    |||
Db 611 CGGTGGACAACTGATCAAGAGAACAAACCTGGCCCTTGTCTGTGGGGACCATTCCTCGGA 552
Oy 2078 GGGACCAAGTTCATGAGAGCCATCACCCTGATGTCAGCAGGGGATGAGATGAGATGAAT 2137
    |||
Db 551 GGGACCAAGTTCATGAGAGCCATCACCCTGATGTCAGCAGGGGATGAGATGAGATGAAT 495
Oy 2138 CGGGGAGAGAGAGGGGCGCCTGCTGCTTGTGACATGACATGACATTCCTGACATGCTTCT 2197
    |||
Db 494 CGGGGAGAGAGAGGGGCGCCTGCTGCTTGTGACATGACATGACATTCCTGACATGCTTCT 435
Oy 2198 GGAAGGTGCTGTTTGGCTGTGCCCCCAGAGAGTACTGCAAGGCTGGGCTGCTTCTG 2257
    |||
Db 434 GGAAGGTGCTGTTTGGCTGTGCCCCCAGAGAGTACTGCAAGGCTGGGCTGCTTCTG 375
Oy 2258 CCGTCTCCATCTCATCATTTGGCAGTCACCGCCATCATTTGGGAGCTGGGCTGCTGCT 2317
    |||
Db 374 CCGTCTCCATCTCATCATTTGGCAGTCACCGCCATCATTTGGGAGCTGGGCTGCTGCT 315
Oy 2318 TGGGGTGCACCAATGCTGTCAAAGATTCAGTCACAGCTGTTTCTGTCGTCATTTGGCA 2377
    |||
Db 314 TGGGGTGCACCAATGCTGTCAAAGATTCAGTCACAGCTGTTTCTGTCGTCATTTGGCA 2355
Oy 2378 CCTGTGTCCAGATAGCTTGGCCAGCAAGCTGCTGCCCTCCAGATGATATGACAGAG 2437
    |||
Db 254 CCTGTGTCCAGATAGCTTGGCCAGCAAGCTGCTGCCCTCCAGATGATATGACAGAG 195
Oy 2438 CCTCCATTTGGCAAGCTGACGAGGAGCAAGCGCTCAATGTCCTTGGGCAATCGGCTG 2497
    |||
Db 194 CCTCCATTTGGCAAGCTGACGAGGAGCAAGCGCTCAATGTCCTTGGGCAATCGGCTG 135
Oy 2498 CCTGTGCTGGGCGCCGATCTACTGGGCTCTGAGGGAAGAGATTCAGAGTCTGGCG 2557
    |||
Db 134 CCTGTGCTGGGCGCCGATCTACTGGGCTCTGAGGGAAGAGATTCAGAGTCTGGCG 75
Oy 2558 GCACACTGGCTTCTCCCTGACCCCTTCTGACATCT-TTGACATTTGTCGATCAGAGTG 2616
    |||
Db 74 GCACACTGGCTTCTCCCTGACCCCTTCTGACATCTCTGACATTTGTCGATCAGAGTG 15
Oy 2617 CTTCTGTACCGAAG 2630
    |||
Db 14 CTTCTGTACCGAAG 1

```

RESULT 3  
B1522813

LOCUS B1522813 792 bp mRNA linear EST 29-AUG-2001

DEFINITION 603175911F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5240066 5',

ACCESSION B1522813

VERSION B1522813.1 GI:15347605

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS NIH-MGC http://mhc.mcl.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LMA11605 row: f column: 03  
High quality sequence stop: 778.  
Location/Qualifiers

```

source 1. 792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5240066"
/clone_11b="NIH_MGC_121"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH-MGC Library."
BASE COUNT 202 a 179 c 227 g 184 t
ORIGIN
Query Match 26.7% Score 739.8; DB 13; Length 792:
Best Local Similarity 98.7%; Pred. No. 1.2e-185;
Matches 767; Conservative 0; Mismatches 7; Indels 3; Gaps 2:
Oy 993 CTATGATCAGCTGGTGGAGTGGCCAAATTAATCTTCTTCCACCAAGAGAGCCG 1052
    |||
Db 1 CTATGATCAGCTGGTGGAGTGGCCAAATTAATCTTCTTCCACCAAGAGAGCCG 60
Oy 1053 CGCCTTCTACCGTATCCAAAGCCATGATGATGACTGTGTCAGGCAATATCTGAGAA 1112
    |||
Db 61 CGCCTTCTACCGTATCCAAAGCCATGATGATGACTGTGTCAGGCAATATCTGAGAA 120
Oy 1113 ACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1172
    |||
Db 121 ACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Oy 1173 TGAGGACTTATTTCCAAAGGCTTCTTGGACCATTTCTTACAGTGGCTGAGAACTG 1232
    |||
Db 181 TGAGGACTTATTTCCAAAGGCTTCTTGGACCATTTCTTACAGTGGCTGAGAACTG 240
Oy 1233 TGGGGCTGTATCTCTGACAGTGTGAGAAAGGGGAGAGACATGTCAAAGCATGTATGT 1292
    |||
Db 241 TGGGGCTGTATCTCTGACAGTGTGAGAAAGGGGAGAGACATGTCAAAGCATGTATGT 300
Oy 1293 GGACTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1352
    |||
Db 301 GGACTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Oy 1353 CAGGCTGTTCTGTAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1412
    |||
Db 361 CAGGCTGTTCTGTAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Oy 1413 CGACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1472
    |||
Db 421 CGACATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Oy 1473 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1532
    |||
Db 481 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Oy 1533 TGTCTAGAGCTCCCTTGTGTGAGCCAGAGTACATCTTGGATGAGAGAGAGAGAGAGAT 1592
    |||
Db 541 TGTCTAGAGCTCCCTTGTGTGAGCCAGAGTACATCTTGGATGAGAGAGAGAGAGAT 600
Oy 1593 CTTCACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1652
    |||
Db 601 CTTCACTTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Oy 1653 GGTTCGCGAGACATCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1711
    |||
Db 661 GGTTCGCGAGACATCAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Oy 1712 GGACAGCCCAAGGCTGGCGGT--GAGGACTTTGAAGACATATAGGAGAGTTGGAATT 1766
    |||

```

| DB                    | 721  | GCAGAGCCAGAGGTGGCGTGCAGCGACTCTGAAACACATATGTGGGAGTTGGAAAT | 777                        |
|-----------------------|--|--|----------------------------|
| RESULT 4              | B0770745   | 775 bp   | RNA linear EST 26-JUL-2002 |
| LOCUS                 | B0770745   |  |                            |
| DEFINITION            | UI-M-F10-BYV-c-07-0-UI.r1 NIH_BMAP_F10   | Mus musculus   | CDNA clone                 |
| IMAGE                 | 6400062 5'   |  | mRNA sequence.             |
| ACCESSION             | B0770745   |  |                            |
| VERSION               | B0770745.1   | GI:21979219  |                            |
| KEYWORDS              | EST.   |  |                            |
| SOURCE                | house mouse.   |  |                            |
| ORGANISM              | Mus musculus   |  |                            |
| REFERENCE             | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |  |                            |
| AUTHORS               | 1 (bases 1 to 775)   |  |                            |
| TITLE                 | NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .  |  |                            |
| JOURNAL               | National Institutes of Health, Mammalian Gene Collection (MGC)   |  |                            |
| COMMENT               | Unpublished (1999)   |  |                            |
|                       | Contact: Robert Strausberg, Ph.D.  |  |                            |
|                       | Email: <a href="mailto:cgabs@emall.nih.gov">cgabs@emall.nih.gov</a>  |  |                            |
|                       | Tissue Procurement: Dr. Jim Lin, University of Iowa  |  |                            |
|                       | CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  |  |                            |
|                       | CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa   |  |                            |
|                       | DNA Sequencing by: Dr. M. Bento Soares, University of Iowa   |  |                            |
|                       | Clone Distribution: MGC clone distribution information can be  |  |                            |
|                       | found through the I.M.A.G.E. Consortium/LLNL at:   |  |                            |
|                       | <a href="http://limage.llnl.gov">http://limage.llnl.gov</a>  |  |                            |
|                       | This clone was contributed by the Brain Molecular Anatomy Project (BMAP)   |  |                            |
| FEATURES              | Seq primer: PYX-5.   |  |                            |
| SOURCE                | Location/Qualifiers  |  |                            |
|                       | 1..775   |  |                            |
|                       | /organism="Mus musculus"   |  |                            |
|                       | /strain="C57BL/6"  |  |                            |
|                       | /db_xref="taxon:10090"   |  |                            |
|                       | /clone="IMAGE:6400062"   |  |                            |
|                       | /clone_lib="NIH_BMAP_F10"  |  |                            |
|                       | /tissue_type="whole brain"   |  |                            |
|                       | /dev_stage="embryo 12.5dpc"  |  |                            |
|                       | /lab_host="DH10B (T1 phage resistant)"   |  |                            |
|                       | /note="Organ: Brain; Vector: PYX-Asc; Site:1: Ecor I; Site:2: Not I; The library was constructed according                           |  |                            |
|                       | Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  |  |                            |
|                       | 1996. Denatured RNA was size fractionated on a 1% agarose  |  |                            |
|                       | gel. First strand cDNA synthesis was primed with oligo-dT  |  |                            |
|                       | primer containing a Not I site. Double strand cDNA was   |  |                            |
|                       | size selected according to mRNA size fraction. ligated   |  |                            |
|                       | with Ecor I adaptor, digested with NotI and then cloned  |  |                            |
|                       | directionally into PYX-Asc vector. The library tag   |  |                            |
|                       | sequence located between the Not I site and the polyA tail   |  |                            |
|                       | is GAGCAGCAGC. This library was created for the University   |  |                            |
|                       | Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  |  |                            |
|                       | Developing Mouse Nervous System', supported by National  |  |                            |
|                       | Institute of Mental Health (NIMH), Hemlin Chlin, Ph.D.,  |  |                            |
|                       | program coordinator."  |  |                            |
| BASE COUNT            | 168 a 185 c 202 g 219 t  | 1  | others                     |
| ORIGIN                |  |  |                            |
| Query Match           | 24.3%  | Score 673.2  | DB 14: Length 775;         |
| Best Local Similarity | 91.7%  | Pred. No. 6.8e-168;                                      |                            |
| Matches 711:          | Conservative 0;  | Mismatches 64;   | Indels 0; Gaps 0;          |
| OY                    | 66 GCTCTTCTGATGATGCTCTTGCAGCAGAGGTGGTGCATCGAGGAGCGTCCAGCAGCAGG   | 125  |                            |
| Db                    | 1 GCTCTTCTGATGATGCTCTTGCAGCAGAGGTGGTGCATCGAGGAGCGTCCAGCAGCAGG  | 60   |                            |
| OY                    | 126 GCAGACATGATGATGCTCTTGCAGGCTCATCGGACTGCAGAGAGGGTGTTCATCTGCCAAT  | 185  |                            |
| Db                    | 61 GCAGACATGATGATGCTCTTGCAGGCTCATCGGACTGCAGAGAGGGTGTTCATCTGCCAAT   | 120  |                            |
| OY                    | 186 CTGCTGACCCGAGAACCTTCCCTTGGGGGACAAATTCAGAGGTCATTTGCTATTTTGT   | 245  |                            |

| Db         | 121   | CTGGATTCACAGAAACCCCTTCCCTTGGGGACAAAGATTGCACGGTCATTTGCTATTTTGT     | 180 |
|------------|---|---|-----|
| QY         | 246   | GACCCTGATPATACATGTTCTTGGGGTGTCCATCATTTGCTGACCGCTTCATGGCATPAT      | 305 |
| Db         | 181   | GGCCCTGATATACATGTTCTTGGGGTGTCTATATTTGCTGCACGATTCATGGCATPAT        | 240 |
| QY         | 306   | TGAATCTATCCACTTCACAAAGAGGGAGGTGACATTTAAGAAACCCANTGGAAACAG         | 365 |
| Db         | 241   | TGAATCTATTCCTTCCACAGAGAGGGAAAGTGACCTTCACAAAGGCCAATGGAGAACAG       | 300 |
| QY         | 366   | CACAAACCACTPATTTGGGTCGTGGATGAAGAAAGTCTCCAAACCTGACCCCTTAATGGCCCTGG | 425 |
| Db         | 301   | CACAACTACANTTGGGGTATGGAAATGAAGAAAGTCTCCAAATCTGACCCGATGAGCCCTGG    | 360 |
| QY         | 426   | TTCTCTGCTCCTCGAGATACCTCTCTCTTTAAATTGAGGTGTGTGTCATGGGTTCAATGGC     | 485 |
| Db         | 361   | CTCTCTCTCCTCCAGAGATTCTCTCTCTTTAAATTGAGGTGTGTGTCACGGGTTCAATGGC     | 420 |
| QY         | 486   | TGGTATCTGGGACCTTTACCATTTGTAGGGAGTGCACCTTCACAACTGTTTCATCATAT       | 545 |
| Db         | 421   | TGGTATCTGGGACCATTTACCATTCGTTGGGAGTGCACCTTCACAACTGTTTCATCATAT      | 480 |
| QY         | 546   | TGGCATCTGTGTCTCTACGTATCCACAGACGAGAGACTGCAAGATCAAGCATCTACAGAGT     | 605 |
| Db         | 481   | TGGCATCTGTGTCTATGTGATCCAGATGGGGAGAGCTGGAAGATCAAGCATCTGAGAGT       | 540 |
| QY         | 606   | CTTCTTCATCAACCGCTGTGGATGATCTTTGCTTACATCTGGGCTTATATGATTTCTGGC      | 665 |
| Db         | 541   | CTTCTTCATCAACCGCTGTGGAGATCTTTGCTTACATCTGGGCTTATATGATTTCTGGC       | 600 |
| QY         | 666   | AGTCTTTCCTCCCTGGTGTGGTCCAGGTTTGGGAAGCCCTCCACTCTCTCTCTCTCTCC       | 725 |
| Db         | 601   | AGTCTTTCCTCCCTGGTGTGGTCCAGGTTTGGGAAGCCCTCTCTCTCTCTCTCTCTCTCC      | 660 |
| QY         | 726   | AGTGTGTCTCCCTTTCGGGCTGGGTGGCAGATTAACGACCTGCTCTTCTAACAATACATGCA    | 785 |
| Db         | 661   | CGTGTGTCTCCCTGCTGGGCTTGGGTGAGATTAAGAGCATGCTCTCTTACAGATACATGCA     | 720 |
| QY         | 786   | CAAAAAGTACCGCACAGACAAAACACCGAGGAATTTATCATAGACACAGGGGTGAC          | 840 |
| Db         | 721   | CAGAAAATACCGCACAGATTAACACCGAGGAATTTATCATAGACACAGGGGTGAC           | 775 |
| RESULT 5   |   |   |     |
| LOCUS      | B0189572  |   |     |
| DEFINITION |   | 665 bp mRNA linear EST 30-APR-2002                                |     |
| ACCESSION  |   | U1-E-EJ1-aka-f-22-0-U1.r1 U1-E-EJ1 Homo sapiens cDNA clone        |     |
| VERSION    | B0189572  |   |     |
| KEYWORDS   | B0189572.1  | GI:20365123   |     |
| SOURCE     | EST.  |   |     |
| ORGANISM   | Homo sapiens  |   |     |
|            | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |   |     |
|            | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.         |   |     |
| REFERENCE  |   |   |     |
| AUTHORS    | Bonaldi,M.F., Lennon,G. and Soares,M.B.                           |   |     |
| TITLE      | Normalization and subtraction: two approaches to facilitate gene  |   |     |
|            | discovery   |   |     |
| JOURNAL    | Genome Res.   | 6 (9), 791-806 (1996)   |     |
| MEDLINE    | 97044477  |   |     |
| COMMENT    |   |   |     |
|            | Contact: Soares, MB   |   |     |
|            | Program for Rat Gene Discovery and Mapping                        |   |     |
|            | University of Iowa  |   |     |
|            | 451 Eckstein Medical Research Building Iowa City, IA 52242, USA   |   |     |
|            | Tel: 319 335 8250   |   |     |
|            | Fax: 319 335 9565   |   |     |
|            | Email: msouares@blue.weeg.uiowa.edu                               |   |     |
|            | Tissue Procurement: Dr. Gregg Hageman                             |   |     |
|            | cDNA Library preparation: Dr. M. Bento Soares, University of Iowa |   |     |
|            | cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  |   |     |
|            | DNA Sequencing by: Dr. M. Bento Soares, University of Iowa        |   |     |
|            | Clone Distribution: Researchers may obtain clones from Research   |   |     |





|   |  |   |                             |
|---|--|---|-----------------------------|
| OY  | 327  | GAGGAGGTGCACATTAAGAACCATAATGGAAACCGACGAACAACCACTATTGGGGCTG    | 386                         |
| Db  | 882  | GGGGGAATCATTCATCAAAAAGCCCAACGGAAACCCACACACACACTAAGGGTTTG      | 823                         |
| OY  | 387  | GAATGAACCTCTCCAACCTGCACCTTATAGGCCCTGGGGTTCCTCCTGCTGAGATACT    | 446                         |
| Db  | 822  | GAAGGAAMCGGTCTCCAACCTCACCCCTCATGGCCCTTGGGGCTGTCCGCCCGAGATCCT  | 763                         |
| OY  | 447  | CCTCTCTTAATTGAAGTGTGTGGTCAATGAGGTTCACTTCTGGTGATCTGGGACCTTAC   | 506                         |
| Db  | 762  | GCTCTCCCTCATCGAGGCTCTCGGGCAGACCTTCAAGCCCGGAGCTGGGAGACCGGCAC   | 703                         |
| OY  | 507  | CATTGTAGGGATGAGAGCTTCAACATCTTCATCTCATTTGGCATCTGTGTCTACGTAT    | 566                         |
| Db  | 702  | CATGCTGGGCGACCGCGGCTTCMACATCTTTGTCTATCTCGGCTGTGCTGTCCGTAT     | 643                         |
| OY  | 567  | CCCAAGACGGAGAGACTCGCAAGATCAAGCATCTACGAGTCTTCTTCATCACCGCTGCTTG | 626                         |
| Db  | 642  | TCCCCAGAGGGGAGGCCCACAGTCAAACA-CTGGGGGTGTCTTCGACAGGGCGGGCTG    | 584                         |
| OY  | 627  | GAGTATCTTTTCCCTACATCTGGCTCTATATGATCTGGCAGCTTCTCCCTGGTGTGT     | 686                         |
| Db  | 583  | GAGACTCTTCTCCCTACATTTGGCTCTACATATCTGTGCGCTGTGTCAGGCCCAAGCTGT  | 524                         |
| OY  | 687  | CCAGGTTTTGGGAAGGCGCTCTACCTCTTCTTCTTTCAGTGTGTGCTTCTGACCTG      | 746                         |
| Db  | 523  | CCAGGTGTGGGAGGGCCM-GTACAGCTGGCTCTCTTCCCGCTGTGCTCTCTGCGCTG     | 465                         |
| OY  | 747  | GGTGGCAGATTAACGACTGCTCTTCTTACAAATACATGACAAAAAGTACCGCACAGCAA   | 806                         |
| Db  | 464  | GTTAGTTCGACCGCGGGCTGTCTTCTTACAACTTCACAAAGATACCGAACCGACAA      | 405                         |
| OY  | 807  | AACCCGAGGAATTTATCATAAGACAGAGGGTGACCAACCTTAAGGSCATTTGATGGATGG  | 866                         |
| Db  | 404  | GCACCGGGGCGGTATCATCGAGACGGAGCGGAGCGCTCGAAGGGGATCTKAGATTGACGG  | 345                         |
| OY  | 867  | GAATAATGATTAATCCACTTTTCTTGAATGGGNAAC-----CTGGTGGCCCT          | 911                         |
| Db  | 344  | CAAGATGTTCAAMTCGCACTTCTGTGATGTGGACGCCGCCCGACACACCTGTGTGGCTGAT | 285                         |
| OY  | 912  | GGAAGGGAAGAGTAGTGAAGTCCCGCAGAGAGATGATCCGGATTTCTCAAGATCTGAA    | 971                         |
| Db  | 284  | GGAGGGCAAAAGAGTGGACRAGTCCCGCCGCAACATGATCCGCATCTTAAAGGATCTGAA  | 225                         |
| OY  | 972  | GCAAAAACACCCAGAGGAAGGACTTGATCAGTGGTGTGAGATGGCCATTTACTATGCTCT  | 1031                        |
| Db  | 224  | GCAGAGCAACCCGGGAGAAAGAGATGACACAGCTGTGTGAGATGGGAAACTACACAGCCCT | 165                         |
| OY  | 1032   | TTC--CCACCAACAGAAAGAGCGCGCTTCTA--CCGTATTCAAAGCCACTCTGATGATGAC | 1088                        |
| Db  | 164  | CTCGTCACACAGAGAAGAGCGCGCTTCTATCTGTGATTCCAAGCAGCGCATATATGAC    | 105                         |
| OY  | 1089   | TGTGTACAGCAATATCTTGAAGAAACATGACAGACAGAACCAAGAAAGCCCTCCAGAT    | 1148                        |
| Db  | 104  | GGGGGCGGGCAACAAATTAAGAAAAAGCAGGTGGGAGGACGACCAAGAGAGCGCAGCGT   | 45                          |
| OY  | 1149   | GAGGAGGTTGCACACCGATGAGCCTGAGGACT                              | 1180                        |
| Db  | 44   | GCAGAGGTTGCACGTGAGAGAGCCCGAAGAGT                              | 13                          |
| <hr/>   |  |   |                             |
| RESULT 8  | BG10325  | 527 bp  | mRNA linear EST 05-JUN-2001 |
| LOCUS   | BG10325  |   |                             |
| DEFINITION  | 60280592JF1 NCL_CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4938316 |   |                             |
| ACCESSION   | BG10325  |   |                             |
| VERSION   | BG10325.1  | GI:14290801   |                             |
| KEYWORDS  | EST.   |   |                             |
| SOURCE  | human.   |   |                             |
| ORGANISM  | Homo sapiens   |   |                             |
| Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |  |   |                             |

| REFERENCE             | AUTHORS   | TITLE  | JOURNAL  | COMMENT  |
|-----------------------|---|--|--|--|
| 1 (bases 1 to 527)    | Mammalia: Eutheria: Primates: Catarrhini; Hominoidea: Homo. | NIH-MGC htp://mgc.nci.nih.gov/.  | National Institutes of Health, Mammalian Gene Collection (MGC) | Unpublished (1999)   |
|                       | Contact: Robert Strausberg, Ph.D.                           | Email: cgapbs-remail.nih.gov   | Tissue Procurement: David N. Louis, M.D.                       |  |
|                       | CDNA Library Preparation: Life Technologies, Inc.           | CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  | DNA sequencing by: Incyte Genomics, Inc.                       | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov |
|                       | Plate: LLM10873   | row: 1   | column: 05   | High quality sequence stop: 527.   |
| FEATURES              | source  | location/Qualifiers  |  |  |
|                       |   | 1..527   |  |  |
|                       |   | /organism="Homo sapiens"   |  |  |
|                       |   | /db_xref="taxon:9606"  |  |  |
|                       |   | /clone_image="4938316"   |  |  |
|                       |   | /clone_id="NCI_CGAP_Brn67"   |  |  |
|                       |   | /tissue_type="anaplastic oligodendroglioma with 1p/19q loss"   |  |  |
|                       |   | /lab_host="DH10B (T1 phage-resistant)"   |  |  |
|                       |   | /note="Organ: Brain; Vector: pCMV-Sport6; Site_1: Not; Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library." |  |  |
| BASE COUNT            | 86 a  | 165 c  | 143 g  | 133 t  |
| ORIGIN                |   |  |  |  |
| Query Match           | 10.7%   | Score 516;   | DB 12;   | Length 527;  |
| Best Local Similarity | 99.8%;  | Pred. No. 4.2e-126;  |  |  |
| Matches 527;          | Conservative 0;   | Mismatches 0;  | Indels 1;  | Gaps 1;  |
| OY                    | 2101  | ACCGCATGTGCGACGAGGGGATGAGATGAGATGAAATCCGGGAGAGAGAGGTGCCCTCC  | 2160   |  |
| Db                    | 1   | ACCGCATGTGCGACGAGGGGATGAGATGAGATGAAATCCGGGAGAGAGAGGTGCCCTCC  | 59   |  |
| OY                    | 2161  | TGCTTTGATACGTATGACACTTCTGTGACTGCTCTGTAAGGTGCTGTTGGCTGTGG   | 2220   |  |
| Db                    | 60  | TGCTTTGATACGTATGACACTTCTGTGACTGCTCTGTAAGGTGCTGTTGGCTGTGG   | 119  |  |
| OY                    | 2221  | CCCCCAGAGACTGTGCGACGCGCTGGGCGTTCCTTCCGCGTCCATCCATCATTTGGC  | 2280   |  |
| Db                    | 120   | CCCCCAGAGACTGTGCGACGCGCTGGGCGTTCCTTCCGCGTCCATCCATCATTTGGC  | 179  |  |
| OY                    | 2281  | ATGCTACCGCCATCATTTGGGGACCTGGCGCTCGCACTTGGCTGCGACCATTTGGTCAAA   | 2340   |  |
| Db                    | 180   | ATGCTACCGCCATCATTTGGGGACCTGGCGCTCGCACTTGGCTGCGACCATTTGGTCAAA   | 239  |  |
| OY                    | 2341  | GATTACGACACGCTGTGTTTCTGGCATTTGGACCTCGTCCGAGATACGTTTCC  | 2400   |  |
| Db                    | 240   | GATTACGACACGCTGTGTTTCTGGCATTTGGACCTCGTCCGAGATACGTTTCC  | 299  |  |
| OY                    | 2401  | AGCAAGGTGCGACGCTCCAGAGATATATATGAGACGCGCTCCATTTGGCAACGTGACGGC   | 2460   |  |
| Db                    | 300   | AGCAAGGTGCGACGCTCCAGAGATATATATGAGACGCGCTCCATTTGGCAACGTGACGGC   | 359  |  |
| OY                    | 2461  | AGCAAGCGCGTCAATGTCTTCTGGGCAATCGGCGCTGGCTGGTCCGTGGCGCGCATTCAC   | 2520   |  |
| Db                    | 360   | AGCAAGCGCGTCAATGTCTTCTGGGCAATCGGCGCTGGCTGGTCCGTGGCGCGCATTCAC   | 419  |  |
| OY                    | 2521  | TGGGCTCTGCGAGGACAGAGATTCACAGTGTGGCGCGGACACATGCGCTTCTCCGTACCC   | 2580   |  |
| Db                    | 420   | TGGGCTCTGCGAGGACAGAGATTCACAGTGTGGCGCGGACACATGCGCTTCTCCGTACCC   | 479  |  |
| OY                    | 2581  | CTCTTCACACATCTTTGCAATTTGTCTGCATACAGCGTGTCTTTTACCGA   | 2628   |  |
| Db                    | 480   | CTCTTCACACATCTTTGCAATTTGTCTGCATACAGCGTGTCTTTTACCGA   | 527  |  |

RESULT 9  
B1756778 718 bp mRNA linear EST 25-SEP-2001  
LOCUS 603024449P1 NIH\_MGC\_114 Homo sapiens cDNA clone IMAGE:5194668 5',  
DEFINITION mRNA sequence.  
ACCESSION B1756778  
VERSION B1756778.1 GI:15748356  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 718)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMNL at:  
http://image.llnl.gov  
Plate: LMNL1487 row: b column: 13  
High quality sequence stop: 644.  
Location/Qualifiers  
1. 718  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5194668"  
/clone\_1id="NIH\_MGC\_114"  
/lab\_host="DH103"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6  
male brains, age range 23-27 yo. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.5 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 019. Note:  
this is a NIH\_MGC Library."  
BASE COUNT 187 a 162 c 219 g 150 t  
ORIGIN  
Query Match 15.1%; Score 416.4; DB 13; Length 718;  
Best Local Similarity 94.8%; Pred. No. 1.8e-99;  
Matches 474; Conservative 0; Mismatches 21; Indels 5; Gaps 4;  
Y 1887 AGATGTGACAGACAGAGAGTGAAGAGAGAGAGAGAGAGATGACAGAGAT 1946  
|||||  
Db 215 AGATGTGACAGACAGAGAGTGAAGAGAGAGAGAGAGAGATGACAGAGAT 274  
Y 1947 GGGAAAGCATTATTTGGTGAACACCCCAACTGAACTCATTTGAAGAGTCCCTATGA 2006  
|||||  
Db 275 GGGAAAGCATTATTTGGTGAACACCCCAACTGAACTCATTTGAAGAGTCCCTATGA 334  
Y 2007 GTTCAAGACTACGGTGAACAACTGATCAAGACAAAGCAAACTGGCTTGGTGGGGAC 2066  
|||||  
Db 335 GTTCAAGACTACGGTGAACAACTGATCAAGACAAAGCAAACTGGCTTGGTGGGGAC 394  
Y 2067 CCATTCTCGAGAGGACAGTTTCATGAGAGCCATCAGTCAGTCAGAGAGGGATGAGGA 2126  
Db 395 CCATTCTCGAGAGGACAGTTTCATGAGAGCCATCAGTCAGTCAGAGAGGGATGAGGA 454  
Y 2127 TGAGAGTGAATCCGGGAGAGAGGAGGCTCCTGCTTTGACTACATGACATTCCT 2186  
|||||  
Db 455 TGAGAGTGAATCCGGGAGAGAGGAGGCTCCTGCTTTGACTACATGACATTCCT 513  
Y 2187 GACTGTCTTCGGAAGGAGGCTTGGCTGTGTGTCCTCCGCCACAGATGATCCACGGCTG 2246  
|||||  
Db 514 GACTGTCTTCGGAAGGAGGCTTGGCTGTGTGTCCTCCGCCACAGATGATCCACGGCTG 573

Y 2247 GGCCTGCTTCGCCCGTTCATCCTCATCATAT--GGCATGCTCACCGCCATTCATTGGGGACC 2305  
|||||  
Db 574 GGCCTGCTTCGCCCGTTCATCCTCATCATATGGGATGTCACCGGCACTATTGGGGACC 633  
Y 2306 TGGCTCGCAGCTT--CGGCTCACCATTTGGTCTCAAGATTCAGTAC--ACCTGTGTTT 2362  
|||||  
Db 634 TGGGCTCGCAGCTTCGGGTCTTAGCCATTGGGCTCAAGATTCAGTACAACTGTGTTGTT 693  
Y 2363 TCGTGCCATTGGGACCGCT 2382  
Db 694 CGGGCAGTTGGACCTCT 713  
RESULT 10  
Bf190598 475 bp mRNA linear EST 02-NOV-2000  
LOCUS Bf190598  
DEFINITION 237175 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.  
ACCESSION Bf190598  
VERSION Bf190598.1 GI:11073967  
KEYWORDS EST.  
SOURCE pig.  
ORGANISM Sus scrofa  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
REFERENCE 1 (bases 1 to 475)  
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,  
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.M.  
and Keeler,J.W.  
Design and use of two pooled tissue normalized cDNA libraries for  
EST discovery in swine  
Unpublished (2000)  
JOURNAL Contact: Smith TPL  
COMMENT USDA, ARS, US Meat Animal Research Center  
PO Box 166, Clay Center, NE 68933-0166, USA  
Tel: 402 762 4366  
Fax: 402 762 4390  
Email: smith@email.marc.usda.gov  
Single pass sequencing. Bases called and alt. trimmed with phred  
v0.980904.e. Vector identified by cross\_match with the -mainscore 18  
and -mismatch 12 options.  
PCR Primers  
FORWARD: AGGAACAGCTATGACCAT  
BACKWARD: GTTTCCCGAGTCACGACG  
Plate: 64 row: H column: 12  
Seq primer: ATTGAGTGACACTATAG.  
FEATURES  
source  
1. 475  
/organism="Sus scrofa"  
/db\_xref="taxon:9823"  
/clone\_1id="MARC 2P1G"  
/tissue\_type="pooled"  
/lab\_host="DH103"  
/note="Vector: pCMV SPORT6; Site\_1: NotI; Site\_2: SalI;  
Library made from pooled tissue from testis, ovary,  
endometrium, hypothalamus, pituitary, and placenta."  
BASE COUNT 68 a 161 c 131 g 115 t  
ORIGIN  
Query Match 14.7%; Score 407.8; DB 12; Length 475;  
Best Local Similarity 91.2%; Pred. No. 2.7e-97;  
Matches 433; Conservative 0; Mismatches 42; Indels 0; Gaps 0;  
Y 2261 TCTCATCTCATCATATTTGGCTGTCACCGGCAATCATTTGGGACCTGGCACTTCG 2320  
|||||  
Db 1 TCTCATCTCATCATATTTGGCTGTCACCGGCAATCATTTGGGACCTGGCACTTCG 60  
Y 2321 GCTCACCATTTGTTCTCAAGATTCAGTCAAGCTGTTGTTTTCGTTGGATTTGGCACT 2380  
|||||  
Db 61 GCTCACCATTTGTTCTCAAGATTCAGTCAAGCTGTTGTTTTCGTTGGATTTGGCACT 120  
Y 2381 CTGTCCAGATAGCTTTGGCAGCAAGCTGCTGCTTCACAGATGATATGACAGCCCT 2440  
|||||



Db 121 CTGTGCCAGATACATTTCGCCAAGACAGCCGCCATCCAGAGCTGTATGACAGCCCT 180  
 QY 2441 CCATTGGCAACGTGACGGGACGACGCGGTCATGTCCTCGGGCATGGGCTGGCCCT 2500  
 Db 181 CCATTGGCAACGTGACGGGACGACGCGGTCATGTCCTCGGGCATGGGCTGGCCCT 240  
 QY 2501 GGTCCGTGGCCGCATCTACTGCGCTTCGAGAGACAGAGTTCACAGTTCGCGCGCA 2560  
 Db 241 GGTCTGTGGCGCCATCTACTGCGCTTCGAGAGACAGAGTTCACAGTTCGCGCGCA 300  
 QY 2561 CACTGCGCTTCCTGTCACCCCTTTCACCATCTTTGCAATTTGTCGATCAGCGTCTCT 2620  
 Db 301 CCTGCGCTTCCTGTCACCCCTTTCACCATCTTTGCAATTTGTCGATCAGCGTCTCT 360  
 QY 2621 TGTACGAGGCGGCGGACGCGGAGGAGGAGCTTGTGCGCCCGCTGGGTGCAAGCTCG 2680  
 Db 361 TGTACGAGGCGGCGGACGCGGAGGAGGAGCTTGTGCGCCCGCTGGGTGCAAGCTCG 420  
 QY 2681 CCACACATGAGCTCTTGTGAGGCTGTGCTCTACATCTCTTGGCACACT 2735  
 Db 421 CCACGACTTGGCTCTTGTGAGGCTGTGCTCTACATCTCTTGGCACACT 475

RESULT 11  
 BH349372/c 499 bp DNA linear GSS 03-DEC-2001  
 LOCUS BH349372  
 DEFINITION CH230-32M17, TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
 ACCESSION BH349372  
 VERSION BH349372.1 GI:17280106  
 KEYWORDS GSS.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus;  
 1 (bases 1 to 499)  
 Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn,  
 A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de  
 Jong, P. and Fraser, C.M.  
 Rat BAC End Sequences from Library CHORI-230 EcORI segment  
 Unpublished (1999)  
 Other\_GSSs: CH230-32M17.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230  
 (http://www.chori.org/bacpac/rat230.htm). For BAC library  
 availability, please contact Pieter de Jong (pje@jengmail.cho.org).  
 Clones may be purchased from BACPAC Resources  
 (http://www.chori.org/bacpac/or.oring.information.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html  
 Plate: 32 row: M column: 17  
 Seq primer: SP6  
 Class: BAC ends.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..499  
 /organism="Rattus norvegicus"  
 /strain="BN/SSNHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-32M17"  
 /clone\_1lb="CHORI-230 Segment 1"  
 /sex="Female"  
 /cell\_type="Brain"  
 /note="Vector: PTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
 CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by  
 Pieter de Jong"

BASE COUNT  
 ORIGIN

Query Match 14.7%; Score 407.4; DB 17; Length 499;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-97;  
 Matches 454; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 361 ACCAGCACAACACATATTCGGGTCTGGAAATGAACGTCTCCAACTGACCTTAT-GGC 419  
 Db 498 ACCAGCACAACATATTCGGGTATGGAATGAACGTCTCCAACTGACCTGATGGCC 439  
 QY 420 CTTGGTCTCTGCT-CCTGAGATACCTCTCTTTAATTTGAGGTGTGTCATGGCT 478  
 Db 438 CATAGGCTTTCCTGCTCCGAGATTCCTCTCTTTAATTTGAGGTGTGTCACGGGT 379  
 QY 479 TCATTCGTGTATCTGAGGACCTTACCATTTTAGGAGATGACCTTAACATGTTC 538  
 Db 378 TCATTCGTGTATTTGGGACCTTACCATTTTGGCAATGACCTTCACATGTTC 319  
 QY 539 TCATCATATGATCTGTGTACGTATCCACAGGAGACCTCCAGATCAAGCATC 598  
 Db 318 TCATCATATGATCTGTGTATGTATGATCCAGATGGGAGACTCCAGATCAAGCACC 259  
 QY 599 TACGAGCTTCTTCATACACCGCTGCTTGAGATCTTTGCTACATCTGGCTATATGA 658  
 Db 258 TTCGAGCTTCTTGTACAGGCTGCTTGAGAGCTTTGCTTATTTGGCTCTACATGA 199  
 QY 659 TTCTGGCAGCTCTCTCCCTGCTGTGTGTCAGATTTGGAGGCTCTCTCTCTCT 718  
 Db 198 TCTTGGCAGCTCTCTCTCTGCTGTGTGTCAGATTTGGAGGCTCTCTCTCTCTCT 139  
 QY 719 TCTTTCAGTGTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 778  
 Db 138 TCTTTCAGTGTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 79  
 QY 779 ACATGCAAAAAGTACCGCAGACACACAGAGAAATATCATAGACAGAGGGTG 838  
 Db 78 ACATGCAAAAAGTACCGCAGACATTAACACAGAGAAATATCATAGACAGAGGGTG 19  
 QY 839 ACCACCTTAAGGCATT 855  
 Db 18 AACACCTTAAGGCATT 2

RESULT 12  
 LOCUS BC023215 3063 bp mRNA linear HTC 04-FEB-2002  
 DEFINITION Homo sapiens, clone IMAGE:4932133, mRNA.  
 ACCESSION BC023215  
 VERSION BC023215.1 GI:18490415  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3063)  
 TITLE Direct Submission  
 AUTHORS Strausberg, R.  
 JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: http://mgc.ncl.nih.gov  
 Contact: MGC help desk  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/  
 Contact: amg@bcm.tmc.edu  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Huilyk, S.W., Hale, S.M.,  
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,



BG819145  
LOCUS 655 bp mRNA linear EST 22-MAY-2001  
DEFINITION 602781252P1 NCI\_CGAP\_Brn67 Homo sapiens CDNA clone IMAGE:4932133  
ACCESSION BG819145  
VERSION BG819145.1 GI:14166732  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10857 row: 9 column: 14  
High quality sequence stop: 653.  
Location/Qualifiers  
1..655  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4932133"  
/clone\_1lb="NCI\_CGAP\_Brn67"  
/tissue\_type="anaplastic oligodendroglioma with 1p/19q  
loss"  
/lab\_host="DH10B (71 phage-resistant)"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site:1; NotI;  
Site:2; SalI; Cloned unidirectionally. Primer: oligo dt.  
Average insert size 2.3 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 156 a 168 c 156 g 175 t  
ORIGIN  
Query Match 13.1%; Score 361.8; DB 12; Length 655;  
Best Local Similarity 97.9%; Pred. No. 5.9e-85;  
Matches 377; Conservative 0; Mismatches 7; Indels 1; Gaps 1;  
QY 2019 GGTGCAAACTGATCAAGAAGCAAAACCTGGCCCTTGGTGGGAGACCATTCTTG-GA 2077  
|||||  
DB 1 GGTGCAAAACTGATCAAGAAGCAAAACCTGGCCCTTGGTGGGAGACCATTCTTGGA 60  
QY 2078 GGGACCACTTCATGAGAGCCATCACCGTCAGTCAGCAGAGGGATGAGATGAGATGAAT 2137  
|||||  
DB 61 GGGACCACTTCATGAGAGCCATCACCGTCAGTCAGCAGAGGGATGAGATGAGATGAAT 120  
QY 2138 CCGGGAGAGAGAGCGCCCTCCCTGCTTGAATGATGATGATGATGATGATGATGATGAT 2197  
|||||  
DB 121 CCGGGAGAGAGAGCGCCCTCCCTGCTTGAATGATGATGATGATGATGATGATGATGAT 180  
QY 2198 GGAAGGTGCTGTGTGCTGTGTCGCCCCACAGAGTACTCCACGGCTGGGCTGTTCG 2257  
|||||  
DB 181 GGAAGGTGCTGTGTGCTGTGTCGCCCCACAGAGTACTCCACGGCTGGGCTGTTCG 240  
QY 2258 CGCTTCATTCATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 2317  
|||||  
DB 241 CGCTTCATTCATCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 300  
QY 2318 TCGGCTGCACATTCGCTGCTGAAGATTCATGACAGCTGTGTTTTCGTCGCAATTGGCA 2377  
|||||  
DB 301 TCGGCTGCACATTCGCTGCTGAAGATTCATGACAGCTGTGTTTTCGTCGCAATTGGCA 360  
QY 2378 CCTCTGTCCAGATACGTTTGGCAG 2402  
|||||  
DB 361 CCTCTGTCCAGATACGTTTGGCAG 385

RESULT 15  
BG293782  
LOCUS 677 bp mRNA linear EST 21-FEB-2001  
DEFINITION 602390677P1 NIH\_MGC\_94 Mus musculus CDNA clone IMAGE:4502515 5',  
ACCESSION BG293782  
VERSION BG293782.1 GI:13053790  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10371 row: 1 column: 20  
High quality sequence start: 14  
High quality sequence stop: 676.  
Location/Qualifiers  
1..677  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4502515"  
/clone\_1lb="NIH\_MGC\_94"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: eye; Vector: PCMV-SPORT6; Site:1; NotI;  
Site:2; SalI; Cloned unidirectionally. oligo-dt primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 176 a 153 c 198 g 150 t  
ORIGIN  
Query Match 13.0%; Score 359.4; DB 12; Length 677;  
Best Local Similarity 93.3%; Pred. No. 2.6e-84;  
Matches 419; Conservative 0; Mismatches 26; Indels 4; Gaps 4;  
QY 1785 GAAACCATTAAGGTTAAATAGTAGTGAAGAGAAATACGAAAGCAAGATTTCTT 1044  
|||||  
DB 227 GAAACCATTAAGGTTAAATAGTAGTGAAGAGAAATACGAAAGCAAGATTTCTT 286  
QY 1845 CATTCGCTTGTGTAACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904  
|||||  
DB 287 CATTCGCTTGTGTAACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 346  
QY 1905 GCTGACTATGGA-AGAAGAGAGAGCCAAAGAGATAGCAGAG-ATGGGAAAGCCAGTATG 1962  
|||||  
DB 347 GCTGACTATGGA-AGAAGAGAGAGCCAAAGAGATAGCAGAG-ATGGGAAAGCCAGTATG 406  
QY 1963 GGTGAACACCCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2021  
|||||  
DB 407 GGTGAACACCCCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 466  
QY 2022 GGCAAACTGATCAAGAAGCAAAACCTGGCCCTTGGTGGGAGCCATTCTCGAGAGGA 2081  
|||||  
DB 467 GGATTAAGCTATCAAGAAGCAAAACCTGGCCATTGGTGGGAGCCATTCTCGAGAGGA 526  
QY 2082 CCAATTGATGAGAGCCATCAACCGTCAGTGCAGAGAGGATGAGATGAGATGATGATGATGAT 2141  
|||||  
DB 527 CCAATTGATGAGAGCCATCAACCGTCAGTGCAGAGAGGATGAGATGAGATGATGATGATGAT 586

Tue Dec 3 11:07:42 2002

us-10-054-680-1.rst

Page 12

| Accession | Sequence   | Position |
|-----------|--|----------|
| Oy        | GGAGGAGAGCGCTGCCCTCCCGCTTACACATGCATGACACTTCGAGACGCTTCTGGAA | 2201     |
|           |  |          |
|           |  |          |
|           |  |          |
| Db        | AGAGGAGAGCGCTGCCACCTCGCTTACACATGCATGACACTTCGAGACGCTTCTGGAA | 646      |
| Oy        | GCTGCTGTTTCCCGTGTGCCCCCAG                                  | 2230     |
|           |  |          |
|           |  |          |
|           |  |          |
| Db        | GCTGCT-CTTGCTGCTGTGCCCCCAGAG                               | 674      |

Search completed: November 30, 2002, 12:23:11  
Job time : 2560 secs